BOUNDED LENGTH CONFIDENCE INTERVALS FOR THE ZERO OF A REGRESSION FUNCTION[1]

By R. H. Farrell

Cornell University

0. Summary. The problem of determining a bounded length confidence interval for the zero of a regression function $R(\cdot)$ is discussed. In case $R(\cdot) = F(\cdot) - p$, F a distribution function, $0 < p < 1$, a closed stopping rule is given for the up-down method of experimentation. For a larger class of regression functions a closed stopping rule is given for Robbins-Monro type of experimentation. The stopping rule for the Robbins-Monro process depends on prior knowledge of an upper and a lower bound on the zero of $R(\cdot)$. It is shown that given suitable assumptions about the random variables used in experimentation finite confidence intervals for the zero of $R(\cdot)$ may be found, such confidence intervals providing an upper and a lower bound on the zero of $R(\cdot)$ with prespecified level of confidence.

1. Introduction. In this paper we discuss the problem of finding bounded length confidence intervals for the zero of a regression function R. We assume there is a number λ such that for all $\theta \,\varepsilon\, (-\infty, \infty)$, if $\theta < \lambda$ then $R(\theta) < 0$, if $\theta > \lambda$ then $R(\theta) > 0$. λ will be called the zero of R (R need not be continuous; we do not assume $R(\lambda) = 0$.)

We will assume given a family of distribution functions $\{G(\cdot, \omega), \omega \,\varepsilon\, \Omega\}$ where Ω is a finite or infinite open real number interval. Certain knowledge of $\{G(\cdot, \omega), \omega \,\varepsilon\, \Omega\}$ is attributed to the experimenter. The exact nature of this knowledge is a mathematical assumption made about a particular technique of experimentation and will be specified in detail in later sections.

In general we will assume that for all $\theta \,\varepsilon\, (-\infty, \infty)$ that $R(\theta) \,\varepsilon\, \Omega$. Also we assume for all $\theta \,\varepsilon\, (-\infty, \infty)$ that $R(\theta) = \int x dG(x, R(\theta))$, this being the justification for calling R a regression function. Last we assume that for all $\theta \,\varepsilon\, (-\infty, \infty)$ the experimenter can observe random variables having $G(\cdot, R(\theta))$ as distribution function.

The statistician's problem is to construct a stochastic process such that given $L > 0$ and $0 < \alpha < 1$, at the termination of experimentation a random interval of length $\leqq L$ has been constructed which covers λ with probability $\geqq 1 - \alpha$. With certain additional assumptions it will be shown how this problem may be solved using a Robbins-Monro process. In the case $\Omega = (0, 1)$ and $\{G(\cdot, \omega), \omega \,\varepsilon\, \Omega\}$ is the family of Bernoulli distribution functions, F a distribution function, $0 < p < 1$, and $R(\theta) = F(\theta) - p$, a closed stopping rule will

Received February 18, 1961.

[1] Research sponsored in part by the Office of Naval Research under Contract No. Nonr-401(03). Part of this material was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Illinois.

be given for an up-down type of method which solves the above problem if F has a unique p-point.

In the construction of bounded length confidence intervals for the zero of R we treat first the problem of finding a finite interval covering the zero of R with a specified confidence level. In Section 3 we show that finite confidence intervals of a given confidence level may be constructed if the following additional assumptions are made:

$$\liminf_{|\theta| \to \infty} |R(\theta)| > 0 \quad \text{and} \quad \{G(\cdot, \omega), \omega \in \Omega\}$$

is a monotone family of distributions (see below.) We suppose $G(\cdot, 0)$ is known to the experimenter. The method to obtain finite confidence intervals uses the properties of certain one sided tests on the parameter ω of the monotone family $\{G(\cdot, \omega), \omega \in \Omega\}$. These tests are characterized by having monotone but discontinuous power functions. Section 2, without reference to the regression problem, gives a description of the construction and some of the properties of these tests.

By a monotone family of distribution functions $\{G(\cdot, \omega), \omega \in \Omega\}$ we mean a family with the property that if $\omega \in \Omega$, $\omega' \in \Omega$, $\omega < \omega'$ then for all $x \in (-\infty, \infty)$, $G(x, \omega) \geqq G(x, \omega')$. It is easily seen that if for some σ-finite measure μ on the Borel sets of the real line each $G(\cdot, \omega)$ is absolutely continuous with respect to μ and if the resulting family of density functions have monotone likelihood ratios then the family of distribution functions $\{G(\cdot, \omega), \omega \in \Omega\}$ is monotone. Consequently in a discussion of the up-down method this hypothesis is automatically satisfied.

In Section 4 it is shown, after additional assumptions are made, how by use of two simultaneous Robbins-Monro processes, a $1 - \alpha$ confidence interval of length $\leqq L$ for the zero of R may be constructed. In Section 5 similar results are given for the up-down method.

In Sections 4 and 5 brief descriptions are given of the two processes. For the reader who is unfamiliar with these stochastic processes we list here some of the previous studies. With the exceptions of Farrell [6] and Tapper (see below) we do not know of other work relating to the subject of this paper.

The Robbins-Monro process is a Markov process which converges to the zero of R. Under differing assumptions various types of convergence may be proven. Convergence in L_2 of the probability space was proven by Robbins and Monro [9]. Almost everywhere point wise convergence has been proven by Blum [1]. Asymptotic normality has been studied by Chung [3], Burkholder [2], and Sacks [10]. For the necessary additional hypotheses the reader should consult the papers mentioned.

The up-down method is a discrete valued Markov process which with probability one reaches every (possible) state infinitely often, as follows from the results of Harris [7]. The use of the up-down method to find LD-50 dosages was studied by Dixon and Mood [5], that is, determine the median of a regression function which is a distribution function F. Dixon and Mood assumed F

is a normal distribution. Derman [4] discusses modifications of the up-down method which allow estimation of a prespecified p-point of F, where F may be an arbitrary distribution function. Mrs. Nancy Tapper, Cornell University, for her doctorial thesis, has been studying closed stopping rules and bounded length confidence interval procedures for the median of F when an estimate is known in advance of the amount of increase in F about its median.

It is assumed in Section 3 (and automatically satisfied in the case of the up-down method) that for each $\omega \varepsilon \Omega$, $\omega = \int x dG(x, \omega)$. If R has jump discontinuities it may happen that $\{G(\cdot, \omega), \omega \varepsilon \Omega\}$ contains distributions which are not observable. In particular if $R(\lambda) \neq 0$ then $G(\cdot, 0)$ is not observable. Yet the inclusion of this distribution is necessary for our analysis of the problem. For this reason our statement of the problem differs somewhat from the usual statement when discussing the Robbins-Monro procedure.

Parts of the material included in this paper appear in the author's thesis, Farrell [6]. The author wishes to thank his advisor, Professor D. L. Burkholder, for his encouragement and for many helpful discussions.

2. One-sided tests. Throughout this section we will assume $\{G(\cdot, \omega), \omega \epsilon \Omega\}$ is a monotone family of distributions in the meaning of Section 1. We will assume Ω is an open real number interval. $\{X_n, n \geqq 1\}$ will be a sequence of mutually independent identically distributed random variables, X_1 having $G(\cdot, \omega)$ as distribution function. For $n \geqq 1$ let $S_n = X_1 + \cdots + X_n$. Suppose $\omega_2 \varepsilon \Omega$ and $\{a_n, n \geqq 1\}$, $\{b_n, n \geqq 1\}$ are extended real number sequences such that if $n \geqq 1$, $a_n > b_n$. Values of $\pm\infty$ are allowed for a_n, b_n. We consider tests of $H_0 : \omega < \omega_2$ against $H_1 : \omega > \omega_2$ having the following form. Let N be the least integer n such that $S_n \geqq a_n$ or $S_n \leqq b_n$, $N = \infty$ if for all $n \geqq 1, b_n < S_n < a_n$. If $S_n \leqq b_n$ accept H_0; if $S_n \geqq a_n$ accept H_1. A subscript "ω" will be used to indicate that $G(\cdot, \omega)$ is the distribution function of X_1; P_ω and E_ω will have corresponding meanings. We will say the test is closed at ω if $P_\omega(N < \infty) = 1$. In general the type of one sided test we will consider here will satisfy $P_\omega(N < \infty) = 1$. In general the type of one sided test we will consider here will satisfy $P_\omega(N < \infty) = 1$ for all $\omega \varepsilon \Omega$, $\omega \neq \omega_2$ and $P_{\omega_2}(N < \infty) < 1$. It will be useful to define four functions of ω, α, β, α_n, β_n by

$$\beta(\omega) = P_\omega(N < \infty, S_N \leqq b_N), \qquad \alpha(\omega) = P_\omega(N < \infty, S_N \geqq a_N),$$

and if $n \geqq 1$,

$$\beta_n(\omega) = P_\omega(N \leqq n, S_N \leqq b_N), \qquad \alpha_n(\omega) = P_\omega(N \leqq n, S_N \geqq a_N).$$

It should be remembered that α, β, α_n, β_n are also functions of the boundaries $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$.

We will now show β, β_n are nonincreasing, α, α_n are non-decreasing, functions of ω. We then show that if $\int x^2 \, dG(x, \omega_2) < \infty$ then, given $\alpha' > 0$, $\beta' > 0$, there exist choices of $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$ such that $\alpha(\omega_2) \leqq \alpha'$ and $\beta(\omega_2) \leqq \beta'$.

Define a function G^{-1} by $G^{-1}(a, \omega) = \inf \{x \mid G(x, \omega) \geqq a\}$. Since G is right continuous, $G^{-1}(a, \omega) \varepsilon \{x \mid G(x, \omega) \geqq a\}$. As is well known, if Z has a uniform distribution on (0, 1) then $G^{-1}(Z, \omega)$ has $G(\cdot, \omega)$ as distribution function. If $\omega < \omega'$ then for all a, $G^{-1}(a, \omega) \leqq G^{-1}(a, \omega')$ as follows at once from the definition.

Let $\{Z_n, n \geqq 1\}$ be a sequence of mutually independent random variables each having a uniform distribution on (0, 1). If $n \geqq 1$ and $\omega \varepsilon \Omega$ we set $S_n(\omega) = \sum_{i=1}^n G^{-1}(Z_i, \omega)$. Let N_ω be the least integer n such that $S_n(\omega) \geqq a_n$ or $S_n(\omega) \leqq b_n$ with $N_\omega = \infty$ if for all $n \geqq 1$, $b_n < S_n(\omega) < a_n$. To prove $\alpha(\cdot)$ is nondecreasing is then to prove $P(N_\omega < \infty, S_N(\omega) \geqq a_N)$ is nondecreasing. We omit "ω" as a double subscript as no ambiguity exists. But if $\omega < \omega'$ then the event $N_\omega < \infty$, $S_N(\omega) \geqq a_n$ implies the event $N_{\omega'} < \infty$, $S_N(\omega') \geqq a_N$. Therefore $\alpha(\cdot)$ is nondecreasing. Similar proofs may be given for the assertions about $\beta(\cdot)$, $\alpha_n(\cdot)$, $\beta_n(\cdot)$.

Now suppose $\int x^2\, dG(x, \omega_2) < \infty$ and $\mu = \int x dG(x, \omega_2)$. As is well known, with probability one, $\lim_{n\to\infty} |S_n - n\mu|/(n^{\frac{1}{2}} \log (n + 1)) = 0$. See for example Loève [8], p. 253, Corollary 2. Therefore with probability one,

$$\sup_{n\geqq 1} |S_n - n\mu|/(n^{\frac{1}{2}} \log (n + 1)) < \infty;$$

for with probability one

$$\sup_{m\geqq n\geqq 1} |S_n - n\mu|/(n^{\frac{1}{2}} \log (n + 1)) < \infty \qquad \text{for every } m \geqq 1.$$

Consequently if $\sup_{n\geqq 1} |S_n - n\mu|/(n^{\frac{1}{2}} \log (n + 1)) = \infty$ on a set of positive measure the same holds for the $\limsup_{n\to\infty}$. We may therefore choose a real number a such that P_{ω_2} (some $n \geqq 1$, $|S_n - n\mu| \geqq an^{\frac{1}{2}} \log (n + 1)) \leqq \min (\alpha', \beta')$. We define $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$ by $a_n = n\mu + an^{\frac{1}{2}} \log (n + 1)$, $b_n = n\mu - an^{\frac{1}{2}} \log (n + 1)$. We have therefore proven the following lemma.

LEMMA. *Suppose Ω is an open real number interval, $\{G(\cdot, \omega), \omega \varepsilon \Omega\}$ a monotone family of distributions. If $\omega_2 \varepsilon \Omega$, $\mu = \int x dG(x, \omega_2)$, $\int x^2\, dG(x, \omega_2) < \infty$ and $\alpha' > 0, \beta' > 0$, then there exist real number sequences $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$ such that if $\omega = \omega_2$, P_ω (some $n \geqq 1$, $S_n \geqq a_n) \leqq \alpha'$, P_ω (some $n \geqq 1$, $S_n \leqq b_n) \leqq \beta'$ and $\lim_{n\to\infty} a_n/n = \lim_{n\to\infty} b_n/n = \mu$. The functions $\alpha(\cdot)$ and $\beta(\cdot)$ are monotone.*

The results obtained above are sufficient for the remainder of this paper. We close this section with a few remarks. For each real number a, $G^{-1}(a, \cdot)$ is a nondecreasing function of its second argument. Therefore if $\omega < \omega'$,

$$\int x\, dG(x, \omega) = EG^{-1}(Z_1, \omega) \leqq EG^{-1}(Z_1, \omega') = \int x\, dG(x, \omega').$$

Equality holds if and only if $G^{-1}(Z_1, \omega) = G^{-1}(Z_1, \omega')$ a.e., that is $G(\cdot, \omega) \equiv G(\cdot, \omega')$. It follows that if for every pair of distinct parameter values the distributions in $\{G(\cdot, \omega), \omega \varepsilon \Omega\}$ are distinct then $\int x dG(x, \omega)$ is a strictly increasing function of ω. In the context of the above lemma, if $\int x dG(x, \omega)$ is a strictly increasing function of ω and the sequences satisfy $\lim_{n\to\infty} a_n/n = \lim b_n/n = \mu$ then if $\omega \neq \omega_2$, $P_\omega(N < \infty) = 1$.

In many examples of monotone families of distributions, given sequences $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$, the functions $\alpha_n(\cdot)$ and $\beta_n(\cdot)$ are continuous in the parameter ω for each $n \geqq 1$. Since $\alpha = \lim_{n\to\infty} \alpha_n$ and $\beta = \lim_{n\to\infty} \beta_n$ it follows that α, β are lower semicontinuous functions. A lower semicontinuous nondecreasing function (α) is always left continuous and a lower semicontinuous nonincreasing function (β) is always right continuous. Given that α_n, β_n are continuous if $n \geqq 1$ Farrell [6] has shown α and β are continuous at each $\omega \in \Omega$ such that $P_\omega(N < \infty) = 1$. He has also shown that if for $n \geqq 1$, α_n and β_n are continuous then α and β can each have at most one point of discontinuity.

If sufficient moments of $G(\cdot, \omega_2)$ are assumed finite, say $\int |x|^3 \, dG(x, \omega_2)$, then sequences $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$ may be constructed using the law of the iterated logarithm.

3. Finite confidence intervals. In the terminology of Section 1 we will assume that the experimenter knows $G(\cdot, 0)$ and that the family $\{G(\cdot, \omega), \omega \in \Omega\}$ is a monotone family of distributions. In constructing tests of the type considered in Section 2 the experimenter need not worry that $\int x^2 \, dG(x, 0) = \infty$. For suppose ϕ is a real valued function defined on $(-\infty, \infty)$ which is strictly increasing such that $|\phi| \leqq 1$. Then the monotone family $\{G(\phi^{-1}(\cdot), \omega), \omega \in \Omega\}$ will serve equally well to determine a finite confidence interval for the zero of R. We will suppose then that $\int x^2 dG(x, 0) < \infty$.

The experimenter may choose real numbers θ and observe random variables having $G(\cdot, R(\theta))$ as their distribution function. If $\{\theta_n, n \geqq 1\}$ is a strictly increasing sequence of real numbers such that $\lim_{n\to\infty} \theta_n = \infty$, then by hypothesis on R, for some n_0, if $n \geqq n_0$ then $R(\theta_n) > 0$. Consequently random variables distributed according to $G(\cdot, R(\theta_n))$ place more probability near $+\infty$ than do random variables distributed according to $G(\cdot, 0)$. We use this idea to construct an upper bound for the zero of R.

THEOREM 1. *Let $\alpha' > 0$ be given and $\{a_n, n \geqq 1\}$ a real number sequence such that P_0 (some $n \geqq 1$, $S_n \geqq a_n$) $\leqq \alpha'$, and $\lim_{n\to\infty} a_n/n = 0$. Suppose $\{\theta_n, n \geqq 1\}$ is a strictly increasing sequence of real numbers such that $\lim_{n\to\infty} \theta_n = \infty$. Let $\{Y_n, n \geqq 1\}$ be a sequence of mutually independent random variables such that Y_n has $G(\cdot, R(\theta_n))$ as distribution function. Let N be the least integer n such that $\sum_{i=1}^n Y_i \geqq a_n$ with $N = \infty$ if for all $n \geqq 1$, $\sum_{i=1}^n Y_i < a_n$. Then if $\liminf_{\theta\to\infty} R(\theta) > 0$, $P(\theta_N \geqq \lambda) \geqq 1 - \alpha'$ and $P(N < \infty) = 1$.*

To prove this theorem we use the ideas of Section 2. We assume λ is the zero of R. Let $\{Z_n, n \geqq 1\}$ be a sequence of mutually independent random variables each uniformly distributed on $(0, 1)$. If $n \geqq 1$ let $Y_n^* = G^{-1}(Z_n, R(\theta_n))$, and $S_n^* = Y_1^* + \cdots + Y_n^*$. Let $0 < \delta = \liminf_{\theta\to\infty} R(\theta)$. There is an $n_1 > 0$ such that if $n \geqq n_1$, $Y_n^* \geqq G^{-1}(Z_n, \delta/2)$. By the strong law of large numbers, with probability one, $\liminf_{n\to\infty} (S_n^*/n) \geqq \delta/2 > 0$. The sequence $\{a_n, n \geqq 1\}$ constructed according to the lemma of Section 2 satisfies $\lim_{n\to\infty} a_n/n = 0$.

Therefore $P(N < \infty) = 1$. Let $\omega = \min(0, R(\theta_1))$

$$P(\theta_N < \lambda) = P(\theta_N < \lambda,\ S_N^* \geqq a_N) \leqq P\left(\theta_N < \lambda,\ \sum_{i=1}^{N} G^{-1}(Z_i, \omega) \geqq a_N\right)$$

$$\leqq P\left(\text{some } n \geqq 1,\ \sum_{i=1}^{n} G^{-1}(Z_i, \omega) \geqq a_n\right) \leqq \alpha'.$$

Therefore $P(\theta_N \geqq \lambda) \geqq 1 - \alpha'$. The proof is complete.

4. Robbins-Monro process. In this section we will show, using the assumptions made below, that confidence intervals of length $\leqq L$ and of confidence level $\geqq 1 - \alpha$ may be constructed for the zero of R. We will suppose the interval length L and the confidence coefficient $1 - \alpha$ given in advance. In order to define certain constants we list the assumptions used before actually defining the Robbins-Monro process.

I. There is a real number λ such that if $\theta < \lambda$ then $R(\theta) < 0$ while if $\theta > \lambda$ then $R(\theta) > 0$.

II. The experimenter can observe random variables X_1 and X_1^* such that $P(X_1^* \leqq \lambda \leqq X_1) \geqq 1 - \alpha/2$.

III. $\sup_{-\infty<\theta<\infty} |R(\theta)|/(1 + |\theta|) = K_1 < \infty$

IV. If $\delta > 0$ then $\inf_{|\theta-\lambda| \geqq \delta} |R(\theta)| > 0$.

V. A real number sequence $\{c_n, n \geqq 1\}$ is given such that $c_n \downarrow 0$,

$$\sum_{n=1}^{\infty} c_n = \infty, \qquad \sum_{n=1}^{\infty} c_n^2 < \infty.$$

VI. There is a function V on Ω defined by

$$V(\omega) = \int (x - \omega)^2\, dG(x, \omega); \qquad \sup_{\omega\varepsilon\Omega} V(\omega) = K_2 < \infty.$$

VII. If $\omega \varepsilon \Omega$, $\omega = \int x\, dG(x, \omega)$.

VIII. For every real θ, $R(\theta) \varepsilon \Omega$.

IX. The functions R and V are Borel measurable.

In order that the results given here be applicable it is assumed that for each real θ the experimenter can observe a random variable $Z(\theta)$ having $G(\cdot, R(\theta))$ as distribution function. In addition we suppose the experimenter knows upper bounds for K_1 and K_2. The results of Sections 2 and 3 give a method by which assumption II may be satisfied. The results of this section do not assume knowledge of $G(\cdot, 0)$ nor that $\{G(\cdot, \omega), \omega \varepsilon \Omega\}$ is a monotone family of distributions.

The Robbins-Monro process is defined as follows. Let X_1 be a random variable. For each $n \geqq 1$ let Z_n be a random variable having conditional distribution function $G(\cdot, R(X_n))$ given $X_1, \cdots, X_n$. Define $X_{n+1} = X_n - c_n Z_n$. It is well known that assumptions weaker than those above imply that with probability one, $\lim_{n\to\infty} X_n = \lambda$. See Blum [1].

In the following we may assume without loss of generality that $\lambda = 0$. For $(X_{n+1} - \lambda) = (X_n - \lambda) - c_n Z_n$ where Z_n has the conditional distribution function $G(\cdot, R((X_n - \lambda) + \lambda)))$. The regression function $R'(\theta) = R(\theta + \lambda)$ has its zero at zero.

THEOREM 2. *Let $A_{n,\delta}$ be the event that for some pair of integers r and s with $n \leqq r < s$, $X_r \leqq 0$ and $X_s \geqq 4\delta$. Assume $4c_n K_1 < 1$ and $c_n K_1 < \delta$. Then*

$$P(A_{n,\delta}) \leqq (K_2/\delta^2) \sum_{i=n}^{\infty} c_i^2.$$

Before proving Theorem 2 we will use it for the construction of confidence intervals. Let X_1 and X_1^* be the random variables specified in II. Suppose c_1 taken so small that $4c_1K_1 < 1$ and $c_1K_1 < \delta$. In addition let the sequence $\{c_n, n \geqq 1\}$ be chosen so that $K_2 \sum_{i=1}^{\infty} c_i^2 < \delta^2(\alpha/4)$. Let $\{X_n, n \geqq 1\}$ be the random variables of a Robbins-Monro process beginning with X_1 and $\{X_m^*, m \geqq 1\}$ the random variables of a process starting with X_1^*. As noted above,

$$0 = \lim_{n\to\infty} X_n = \lim_{m\to\infty} X_m^*$$

with probability one. Let $\epsilon > 0$ be given and N, M respectively be the least integers n, m such that $|X_n - X_m^*| \leqq \epsilon$. We show that with probability $\geqq 1 - \alpha$ the interval $(-4\delta + X_M^*, 4\delta + X_N)$ contains 0 (and hence is nondegenerate). This is an interval of length $\leqq 8\delta + \epsilon$. Observe that

$$\begin{aligned} P(\text{fail to cover } 0) &\leqq P(X_1 < 0 \text{ or } X_1^* > 0) \\ &\quad + P(X_1 \geqq 0, \text{ some } n \geqq 1, X_n + 4\delta < 0) \\ &\quad + P(X_1^* \leqq 0, \text{ some } m \geqq 1, X_m^* - 4\delta > 0) \leqq \alpha/2 + \alpha/4 + \alpha/4 = \alpha \end{aligned}$$

by virtue of assumption II and Theorem 2.

In order to prove Theorem 2 the arguments which prove almost everywhere convergence are reexamined. Define for $n \geqq 1$, $u_n = X_{n+1} - X_n + c_n R(X_n)$. Then if $n \geqq 1$, $u_n = c_n(R(X_n) - Z_n)$ and $E(u_n/X_1, \cdots, X_n) = 0$.

$$E(u_n^2) \leqq c_n^2 K_2.$$

The sequence of sums $\{\sum_{j=1}^{n} u_j, n \geqq 1\}$ is a martingale. From the semimartingale inequality it follows that for

$$\delta > 0, P\left(\max_{m \leqq j \leqq n} \left|\sum_{i=m}^{j} u_i\right| > \delta\right) \leqq (K_2/\delta^2) \sum_{i=m}^{n} c_i^2.$$

Suppose $P(A_{n,\delta}) > (K_2/\delta^2) \sum_{i=n}^{\infty} c_i^2$. Then with positive probability there exists a pair of integers r and s with $n \leqq r < s$ such that $X_r \leqq 0$ (assume $\lambda = 0$), $X_i > 0$ for $i = r + 1, \cdots, s$, $X_s \geqq 4\delta$, $|\sum_{i=r}^{s-1} c_i(R(X_i) - Z_i)| \leqq 2\delta$. Since $X_s - X_r + \sum_{i=r}^{s-1} c_i R(X_i) = \sum_{i=r}^{s-1} c_i(R(X_i) - Z_i)$ we have that

$$X_s - X_r \leqq 2\delta - c_r R(X_r),$$

since $\sum_{i=r+1}^{s-1} c_i R(X_i) \geqq 0$. Therefore $X_s - X_r \leqq 2\delta + c_r K_1(1 + |X_r|) \leqq 2\delta + c_r K_1(1 + X_s - X_r - 4\delta)$. Solving for $X_s - X_r$ yields

$$X_s - X_r \leqq (2\delta + c_r K_1(1 - 4\delta))/(1 - c_r K_1) < 3\delta(\tfrac{4}{3}) = 4\delta.$$

Contradiction. Therefore $P(A_{n,\delta}) \leqq (K_2/\delta^2) \sum_{i=n}^{\infty} c_i^c$.

One type of regression function considered by Robbins and Monro [9] was $R(\theta) = F(\theta) - p$, F a distribution function, $0 < p < 1$. For each $\theta \,\varepsilon\, (-\infty, \infty)$, $P(Z(\theta) = 1 - p) = F(\theta)$, $P(Z(\theta) = -p) = 1 - F(\theta)$. Then

$$K_1 = \sup_\theta |F(\theta) - p|/(1 + |\theta|) \leqq \max(1 - p, p)$$

and $K_2 \leqq \sup_\theta F(\theta)(1 - F(\theta)) \leqq \frac{1}{4}$. The restrictions on c_1 are $4c_1 \leqq 1$ and $c_1 < \delta$. Then $P(A_{n,\delta}) \leqq (\frac{1}{4}\delta^2) \sum_{i=n}^{\infty} c_1^2$.

5. The up-down method. One technique of finding the 50 per cent lethal dosage of a drug is the up-down method. The essential features of this method may be described as follows. For each real number θ, $F(\theta)$ is the probability of "killing" if the "dose" θ is used. A sequence $\{\theta_n, -\infty < n < \infty\}$ is chosen such that for every integer n, $\theta_{n+1} > \theta_n$ and $\lim_{|n|\to\infty} |\theta_n| = \infty$. The initial experiment is made at level θ_0. If the nth experiment is made at level $\theta_{N(n)}$ then the next experiment is made at level $\theta_{N(n+1)}$ where $N(n + 1) = N(n) + 1$ if the individual lives, $N(n + 1) = N(n) - 1$ if the individual dies.

A model for this process may be constructed as follows. Let a function g of two real variables be defined as follows: $g(x, y) = -1$ if $x \leqq y$ and $g(x, y) = 1$ if $x > y$. Let $\{Z_n, n \geqq 0\}$ be a sequence of mutually independent random variables each uniformly distributed on $(0, 1)$. Define $N(0) = 0$ and if $n \geqq 0$, $N(n + 1) = N(n) + g(Z_n, F(\theta_{N(n)}))$.

The possible states for this Markov process are the integers m such that $0 < F(\theta_m) < 1$ together with the greatest integer m satisfying $F(\theta_m) = 1$, $F(\theta_{m-1}) < 1$ and the least integer m satisfying $F(\theta_m) = 0$, $F(\theta_{m+1}) > 0$. Let A be this set of integers. If $0 \not\varepsilon A$ it is clear that the process changes monotonely until a value in A is reached. From that point, with probability one, all subsequent values lie in A. We assume in the sequel that $0 \,\varepsilon\, A$.

It follows at once from the results of Harris [7] that if $i \,\varepsilon\, A$ then with probability one $N(n) = i$ infinitely often. Define an integer valued random variable $M(i, n)$ to be the least integer m such that $N(0), \cdots, N(m)$ takes the value i exactly n times. A direct calculation shows $\{g(Z_{M(i,n)}, F(\theta_i)), n \geqq 1\}$ is a sequence of independent identically distributed random variables such that $P(g(Z_{M(i,1)}, F(\theta_i)) = -1) = F(\theta_i)$. If $i \,\varepsilon\, A$ then with probability one this sequence is infinite.

We turn to the construction of confidence intervals for the p-point λ_p of F. We suppose if $\theta < \lambda_p$ then $F(\theta) < p$ and if $\theta > \lambda_p$ then $F(\theta) > p$. Let

$$\{Y_n, n \geqq 1\}$$

be a sequence of independent identically distributed Bernoulli random variables such that $P(Y_1 = 1) = q$. We introduce these random variables only as a device to describe integer sequences $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$. According to the lemma of Section 2 we may choose integer sequences $\{a_n, n \geqq 1\}$ and $\{b_n, n \geqq 1\}$ such that if $q \leqq p$, $P(\text{some } n \geqq 1, \sum_{i=1}^n Y_i \geqq a_n) \leqq \alpha/4$ and if $q \geqq p$ then $P(\text{some } n \geqq 1, \sum_{i=1}^n Y_i \leqq b_n) \leqq \alpha/4$. Further $\lim_{n\to\infty} a_n/n = \lim_{n\to\infty} b_n/n = p$ is assumed for these sequences. $1 - \alpha$ is the desired confidence level.

We now show that while up-down experimentation continues random integers I and J are determined such that $P(\theta_I \leqq \lambda_p) \geqq 1 - \alpha/4$ and $P(\theta_J \geqq \lambda_p) \geqq 1 - \alpha/4$. Define random integer sequences $\{c(n), n \geqq 1\}$ and $\{d(n), n \geqq 1\}$ as follows.

(1) $c(1) = d(1) = 0$.

(2) if $m \geqq 1$, $c(m + 1)$ is the least integer n such that $N(n) \geqq N(c(m))$ and $n > c(m)$.

(3) if $m \geqq 1$, $d(m + 1)$ is the least integer n such that $N(n) \leqq N(d(m))$ and $n > d(m)$.

It will be seen that $N(c(1)), \cdots, N(c(n)), \cdots$ is the sequence of highest levels of experimentation reached while $N(d(1)), \cdots, N(d(n)), \cdots$ is the sequence of lowest levels reached.

Since each $i \varepsilon A$ is reached infinitely often (with probability one), with probability one, if $i \varepsilon A$, for some $n \geqq 1$, $N(c(n)) \geqq i$ and with probability one for some $n \geqq 1$, $N(d(n)) \leqq i$. Define integer valued random variables $\bar{M}$ and $\underline{M}$. $\bar{M}$ is the least integer n such that

$$(\tfrac{1}{2}) \sum_{i=1}^{n} (1 - g(Z_{N(c(i))}, F(\theta_{N(c(i))}))) \geqq a_n .$$

$\underline{M}$ is the least integer n such that

$$(\tfrac{1}{2}) \sum_{i=1}^{n} (1 - g(Z_{N(d(i))}, F(\theta_{N(d(i))}))) \leqq b_n .$$

To understand these inequalities recall g takes ± 1 as its values so that $(\frac{1}{2})(1 - g)$ takes values 0, 1. We suppose $\bar{M} = \infty$ or $\underline{M} = \infty$ if the required inequality never holds but show at once $P(\bar{M} < \infty, \underline{M} < \infty) = 1$. Assume for definiteness that $\theta_{-1} < \lambda_p \leqq \theta_0$. Then $F(\theta_{-1}) < p < F(\theta_1)$ and $-1 \varepsilon A$, $1 \varepsilon A$. With probability one, $N(c(i))$ eventually reaches values $\geqq 1$ and $N(d(i))$ eventually reaches values $\leqq -1$. It follows that with probability one

$$(\tfrac{1}{2})(1 - g(Z_{N(c(i))}, F(\theta_{N(c(i))})))$$

eventually becomes $\geqq (\frac{1}{2})(1 - g(Z_{N(c(i))}, F(\theta_1)))$.

$$\{(\tfrac{1}{2})(1 - g(Z_{N(c(i))}, F(\theta_1))), i \geqq 1\}$$

is a sequence of independent identically distributed Bernoulli random variables such that

$$P((\tfrac{1}{2})(1 - g(Z_{N(c(1))}, F(\theta_1))) = 1) = F(\theta_1).$$

By the strong law of large numbers and using the assumption $\lim_{n\to\infty} a_n/n = p$ it follows that $P(\bar{M} < \infty) = 1$. Similarly $P(\underline{M} < \infty) = 1$.

By an argument similar to that used to prove Theorem 1, if $\theta_{N(c(\bar{M}))} < \lambda_p$ then $N(c(i)) \leqq -1$ for $i = 1, \cdots, \bar{M}$ so that $P(\theta_{N(c(\bar{M}))} < \lambda_p) \leqq \alpha/4$. Similarly if $\theta_{N(d(\underline{M}))} > \lambda_p$ then $N(d(i)) \geqq 0$ for $i = 1, \cdots, \underline{M}$ and $P(\theta_{N(d(\underline{M}))} > \lambda_p) \leqq \alpha/4$.

Having determined integers $I = N(d(\underline{M})) < J = N(c(\bar{M}))$ experimentation is continued at the levels $\{\theta_i, I < i < J\}$. Note that if n is the number of the observation at which the values of $\bar{M}$ and $\underline{M}$ are decided then $N(n) = I$ or $N(n) = J$. Restriction of experimentation to the interval $I < i < J$ means that whenever subsequently the value $N(n) = I$ or $N(n) = J$ is reached some rule is used to choose a value i, $I < i < J$ for the next experiment. Otherwise the method is not changed.

Experimentation is continued until the following conditions are satisfied. Let P be the total number of observations taken.

(4) If $I < i < J$, $P(i)$ is the number of $N(0), \cdots, N(P)$ taking the value i. Note that observations taken prior to deciding the values I, J are included.

(5) If $I < i < J$, except possibly one value of i in this range,

$$(\tfrac{1}{2}) \sum_{j=1}^{p(i)} (1 - g(Z_{M(i,j)}, F(\theta_i))) \geqq a_{P(i)} \quad \text{or} \quad \leqq b_{P(i)} .$$

That is, the outcomes of experiments at level θ_i are examined for the frequency of deaths. This frequency is compared against the sequences $\{a_n, n \geqq 1\}$ or $\{b_n, n \geqq 1\}$. The condition for stopping includes two consistency conditions stated next.

(6) If $I < i < k < J$ and $(\frac{1}{2}) \sum_{j=1}^{P(i)} (1 - g(Z_{M(i,j)}, F(\theta_i))) \geqq a_{P(i)}$, then $(\frac{1}{2}) \sum_{j=1}^{P(k)} (1 - g(Z_{M(k,j)}, F(\theta_k))) \geqq a_{P(k)}$.

(7) If $I < k < i < J$ and $(\frac{1}{2}) \sum_{j=1}^{P(i)} (1 - g(Z_{M(i,j)}, F(\theta_i))) \leqq b_{P(i)}$, then $(\frac{1}{2}) \sum_{j=1}^{P(k)} (1 - g(Z_{M(k,j)}, F(\theta_k))) \leqq b_{P(k)}$. Let I^* be the greatest integer i such that $I \leqq i < J$ and

$$(\tfrac{1}{2}) \sum_{j=1}^{P(i)} (1 - g(Z_{M(i,j)}, F(\theta_i))) \leqq b_{P(i)} .$$

Let J^* be the least integer i such that $I < i \leqq J$ and

$$(\tfrac{1}{2}) \sum_{j=1}^{P(i)} (1 - g(Z_{M(i,j)}, F(\theta_i))) \geqq a_{P(i)} .$$

Then $P(\theta_{I^*} \leqq \lambda_p \leqq \theta_{J^*}) \geqq 1 - \alpha$ and $P(J^* - I^* \leqq 2) = 1$.

We have assumed that $\theta_{-1} < \lambda_p \leqq \theta_0$. Observe that

$$\begin{aligned} 1 - P(\theta_{I^*} \leqq \lambda_p \leqq \theta_{J^*}) &\leqq P(\theta_I > \lambda_p) + P(\theta_J < \lambda_p) \\ &+ P((\tfrac{1}{2}) \sum_{j=1}^{P(-1)} (1 - g(Z_{M(-1,j)}, F(\theta_{-1}))) \geqq a_{P(-1)}) \\ &+ P((\tfrac{1}{2}) \sum_{j=1}^{P(0)} (1 - g(Z_{M(0,j)}, F(\theta_0))) \leqq b_{P(0)}) \leqq 4(\alpha/4) = \alpha. \end{aligned}$$

That $J^* - I^* \leqq 2$ follows from conditions (4) to (7). That these conditions will be eventually satisfied (with probability one) follows from the strong law of large numbers together with the hypothesis $\lim_{n\to\infty} a_n/n = \lim_{n\to\infty} b_n/n = p$.

REMARKS. We have given a confidence interval procedure for the up-down method which retains the essential properties of the method. It should be observed, however, that the only properties of the up-down method that ultimately enter the probability analysis are that observations are taken only at levels $\{\theta_n, -\infty < n < \infty\}$ and that each possible state $i \in A$ is reached infinitely often with probability one. A criticism sometimes made of the up-down method is that the experimenter is required to take one observation at each step. Although the procedure described has this characteristic it is not inherent in our procedure. The probability statements made depend only on the comparisions of sums of independent Bernoulli random variables against the integer sequences

$$\{a_n, n \geqq 1\} \qquad \text{and} \qquad \{b_n, n \geqq 1\}.$$

No matter how an integer valued random variable Q is defined, for example, $P_q(Y_1 + \cdots + Y_Q \geqq a_Q) \leqq P_q(\text{some } n \geqq 1, Y_1 + \cdots + Y_n \geqq a_n) \leqq \alpha/4$ if $q \leqq p$. The notations are as defined earlier.

The experimenter has the option of defining a procedure as best suits his needs. The requirements are that he should obtain bounds $I < J$ satisfying $P(\theta_I \leqq \lambda_p \leqq \theta_J) \geqq 1 - \alpha/2$ and that sufficient experimentation be made at the levels $i, I < i < J$ to obtain a consistent set of inequalities $\theta_i \geqq \lambda_p$ or $\theta_i \leqq \lambda_p$ for all except possibly one $i \in (I, J)$; this consistency is the meaning of requirements (6) and (7) above. The levels of experimentation may be chosen by any rule not depending on the future and consequently any number of observations may be taken at a given level before moving to a different level. The only requirement is that a sum S based on k observations is to be compared with a_k or b_k when making decisions to stop or continue.

REFERENCES

[1] J. R. BLUM, "Approximation methods which converge with probability one," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 382–386.

[2] D. L. BURKHOLDER, "On a class of stochastic approximation procedures," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1044–1059.

[3] K. L. CHUNG, "On a stochastic approximation method," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 463–483.

[4] C. DERMAN, "Non-parametric up-and-down experimentation," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 795–798.

[5] W. J. DIXON AND A. M. MOOD, "A method for obtaining and analyzing sensitivity data," *J. Amer. Stat. Assn.*, Vol. 43 (1948), pp. 109–126.

[6] R. H. FARRELL, "Sequentially determined bounded length confidence intervals," Thesis, University of Illinois, 1959.

[7] T. E. HARRIS, "First passage and recurrence distributions," *Trans. Amer. Math. Soc.*, Vol. 73 (1952), pp. 471–486.

[8] MICHEL LOÈVE, *Probability Theory*, D. Van Nostrand, New York, 1955.

[9] H. ROBBINS AND S. MONRO, "A stochastic approximation method," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 400–407.

[10] JEROME SACKS, "Asymptotic distribution of stochastic approximation procedures," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 373–405.